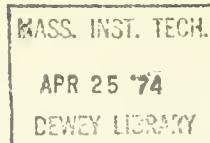




APR 9 75

3
14
-74

Dewey



WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

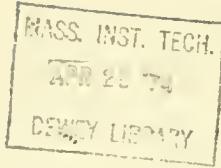
TOWARDS A BEHAVIORAL METHODOLOGY
FOR THE STUDY OF OR/MS
IMPLEMENTATION

P.G.W. KEEN
SLOAN SCHOOL OF MGT.
MIT

FEBRUARY 1974

701-74

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139



TOWARDS A BEHAVIORAL METHODOLOGY
FOR THE STUDY OF OR/MS
IMPLEMENTATION

P.G.W. KEEN
SLOAN SCHOOL OF MGT.
MIT

FEBRUARY 1974

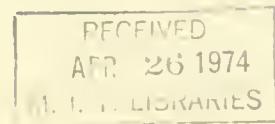
701-74

This is a preliminary version; comments and criticisms are requested

HD28

.M414

no. 701-74



This paper examines some methodological issues in behavioral research focussed on implementing OR and MS applications. This research is becoming increasingly important within the OR/MS field and at present is mainly concerned with identifying the contingent factors or organization, personality or context that facilitate implementation. It is thus exploratory and lacks firm theoretical paradigms or strategies for research. The paper discusses limitations of the common 'Factor Study' which uses attitude surveys or opinion questionnaires and inductive multivariate analysis to derive a mapping of the relevant contingencies. A strategy for this type of behavioral research is outlined: it stresses the importance of examining implementation as a process and not in terms of structural variables, and the need for a complete initial field description as the base for generalizing from the conclusions of the detailed data gathered in the research effort. Guidelines for the selection of data and of instruments to measure it are presented; the paper criticizes the use of surrogate variables based on convenience and strongly suggests a diagnostic approach based on taxonomic classification of the main contingencies, reinforced by a 'rhetorical', supportive selection of both what data to collect, the instruments for doing so, and the type of formal analysis used to evaluate the data. The aim of this type of research should be to provide 'generative' conclusions for theories of the middle range (Merton).

The paper is intended for two audiences: OR/MS academics and practitioners who have begun to extend their focus from the technical to the behavioral and researchers studying behavior in organizations whose interests run parallel to, but do not include, the study of complex computer-based projects. Parts of the material will be self-evident to one or the other of these two audiences.

It is only very recently that the focus[✓] of attention in studies of OR/MS implementation has shifted from the technical to the behavioral. The change has been swift and seems to indicate a level of maturation within the field of both OR/MS research and practice; there is no longer a need for theological debates about, for instance, simulation versus optimization, and on the whole, one may view any particular technique with the "willing suspension of disbelief" advocated by Coleridge for the reading of poetry. The earlier focus[✓], which was specifically analytic and technical, was marked by a high conceptual level, rigorous definitions and a distinctively normative zeal. The behavioral approach is, by contrast, more descriptive, lacking clear conceptual frameworks and - at this point at least - very uncertain in its terms and methods. It has a wider scope, the study of OR/MS within the organization as a whole, and because of this must explicitly examine a wider range of interrelated variables. Many behavioral studies of implementation aim merely at clarifying those variables. It is apparent that in such studies the conceptual problems are few, but the methodological issues immense. The simplest solution to those issues has been to borrow from the grab bag of experimental techniques available in the social sciences, particularly social psychology. The criterion for making this choice is generally convenience rather than suitability. This has resulted in many behavioral studies of implementation examining complex organizational phenomena in terms of the variables most accessible or easily measured and exploring systematic interrelations through particular multivariate techniques, especially regression and factor analysis. This approach is certainly convenient; any alternative strategy immediately encounters substantial problems. However, it is surely essential that studies of implementation start from a clear conceptual perspective and tie their methodology into that base. The relationship between theory and method is an intimate one; it is apparent that several recent studies constrain their

theoretical schemes by their borrowed methodology. They show an underlying axiom: "if we can't measure it it's not important and if we do measure it it is". It is always easy to point to inadequacies in current practice rather than to suggest better approaches. The aim of this paper is to provide, with reasonable specificity, a conceptual and methodological base for the study of implementation; its starting point is, none the less, the mistakes of others. The inadequacies of some recent studies does seem to merit a critique in itself, but more importantly it is informative to look at the reasons researchers have for choosing a methodology that is distinctly flawed, focusing on what they obtain from that method and what they aim at achieving.

There is a single simple question that underlies most behavioral studies of implementation: "What factors influence the likelihood of success in an OR/MS effort?" We have very little conventional wisdom about the contingencies affecting implementation; it seems unlikely that there are many absolutes, although some approaches, such as early studies of decentralization versus centralization, seem to be based on the assumption that there are. Most behavioral studies of OR/MS aim at mapping contingencies of organizational structure, strategy for implementation, management style, communication, etc., etc., starting essentially from a position of ignorance. This approach may reasonably be termed the "Factor Study". It focuses at a very general level on the overall issue of implementation. The intended output of the study is a better map.

There is a distinctive variant of the Factor study that will serve very well as the straw man around which to structure this paper. It is fairly common, but it seems of little value to cite specific instances here merely to chop them down. The discussion below is based on several recent papers and uses disguised examples derived from them. The format of this style of research is generally a questionnaire, or more grandiloquently a "Likert survey"; this is a series of statements to which the respondent indicates the extent to which he agrees

through a 5 or 7-point Likert scale. (The scale is scored on an ordinal basis, with a "1" corresponding to "strongly disagree" and a 5 or 7, whichever is the highest value for the chosen scale-indicating "strongly agree"). Exhibit 1 shows sample statements (out of a total of 100) from a questionnaire administered to 80 managers in a large organization identified as a "successful" user of OR/MS. The responses are scored and correlated (product-moment coefficients) and the correlation matrix used as input to a factor analysis program. Factor analysis essentially groups the responses, structuring the correlations into a set of "factors"; these are unnamed and the researcher must attach interpretation and labels. The decision rules for factor analysis are complex, but the most commonly used methods derive a "simple structure" that best explains the variance among the intercorrelations. Exhibit 2 shows the results for the questionnaire in the example of Exhibit 1.

In principle, the factors shown in Exhibit 2 place a dynamic structure on a systemic phenomenon. Out of the forest of the 100 statements, the analysis has picked the seven main trees. Factor analysis is a powerful, if controversial technique, with some extremely valuable applications. The criticisms of this paper relate to the factor study and not to factor analysis itself. It is worth mentioning, however, J. Scott Armstrong's delightful experiment in which he generated meaningful factors via random numbers ("Derivation of Theory by Means of Factor Analysis or Tom Swift and his Electric Factor Analysis Machine").¹ The results of Exhibit 2 imply that there is a dominating factor, indicated by the commonality among the responses to statements 6, 19, 55, and 71 on the questionnaire; from examining the statements, the researcher identifies this factor as "the analyst's technical competence". The second most dominant factor is similarly categorized as "The immediacy of the problem". Together these two factors account for 64% of the total variance among the correlations. By contrast, the seventh factor, identified as "analyst/client

EXHIBIT 1 Sample "Likert" Questions

The PLANIT model will help me save time looking for information.

The complexity of the decisions I make requires the use of quantitative methods.

On the whole, our OR analysts communicate ideas well.

The PLANIT project is technically sound

OR personnel need to have management experience

PLANIT will help me meet my performance goals

EXHIBIT 2 Factors Derived from Questionnaire Results

Factor 1 "Analyst's technical competence"

(Factor length .36)

Loadings Question

.59 6 The design of the model should be the OR group's responsibility

.52 19 Our OR group is exceptionally capable

-.47 55 OR personnel need to have management experience

.46 71 I generally have confidence in the recommendations the OR group make

Factor 2 "Immediacy of the problem"

(Factor length .28)

.81 18 We need PLANIT now

.61 64 The PLANIT project is important to my department

.48 49 I intend to use PLANIT

-.39 2 The development costs for PLANIT have been too high

Factor 3 "Top management support"

.

.

.

.

Factor 7 "Client/analyst interface"

(Factor length .03)

interface", accounts for just 3%. From the full analysis, the researcher may (and in the example does) induct a set of general conclusions. From the "explanation" of the whole questionnaire implied in the factor structure, it would seem that the main strategy for implementation should be to develop confidence in the OR/MS group's technical skill rather than to build "trust" and also to take on projects only where management needs a quick answer to a pressing problem. This is what the map provided by the analysis clearly suggest.

The strength of the Factor study - conceptually - is that it imposes a frame on the system it examines as a whole and does not explore it in any reductionist manner isolating single predictive variables. It takes a complex, dynamic, blurred situation and derives meaning from it. It also does justice to the multivariate nature of its subject. Not least, moreover, the techniques of factor analysis are highly convenient particularly if a time-shared program is available; it is fair to suggest that no other technique massages so much data so quickly, with such immediately usable output. Since the main problem in behavioral research on implementation is to uncover the molar clusters of variables of most impact on the whole situation, the inductive data-reduction, data-structuring approach of the factor study has obvious appeal.

Regardless of its appeal, however, this style of study is essentially invalid in all respects. Most critically, it ignores contingencies and inducts generalization from highly specific contexts. The example shown in Exhibit 2 partly derives from a conference paper that argued that the designer/client interface, particularly in relation to interpersonal features, was of little importance to the success of modelling projects. The paper did, however, include the name of the organization concerned, a division of a large chemical company. It seems reasonable to presume that the managers of that division would tend to hold degrees in technical or scientific subjects. It is clearly apparent that the technology involved in the manufacture of a wide range of chemical

products from a limited variety of raw materials is well-suited to optimization techniques. The division has used OR/MS for some years and has a well-established MS group with a formal charter. Given these contingent factors, it seems hardly surprising that the organization has few problems with the designer/client interface; compatibility of manager and analyst is provided through a similarity of training and probably of outlook, the problems are well-suited to direct input to quantitative methods and there is a background of shared experience and association. The Factor study in this example does not recognize these contingencies; the lack of any prior conceptual scheme makes it, perhaps, impossible that it could do so - the analysis of the particular data gathered through the questionnaire derives a plausible model without any criteria for assessing the outcome, except through repeating the experiment with the derived scheme used as an a priori focus. It is ironic that the absence of a factor is in no way an indication of its lack of importance; in the example the designer/client interface may well be vital but not apparent, because it is uniformly regarded by the managers as satisfactorily handled (the result will be a narrow range of scores on the relevant statements for the Likert scales - this lack of spread will affect the correlation matrix, the bedrock for the study, by failing to show relationships that may well exist). In addition, the factors derived obviously depend on what questions are asked in the Likert scales.

The second weakness of this style of Factor study is more subtle. The Likert scale questionnaire samples attitudes and opinions. This obviously assumes that attitudes are adequate indicators of action, that there is an association between the behavior resulting from an OR/MS effort and the attitudes of managers towards the effort. Social psychology for many years operated on the basis of that axiom; in fact the concept of attitude in itself opened up research in that field. However, the axiom has been found to be

dubious at best. Festinger, in 1963, describes his own reaction to a suggestion that, perhaps, attitude is not a good predictor of behavior:

"I was, at first reading, slightly skeptical about the assertion that there is a dearth of studies relating attitude or opinion change to behavior. Although I could not think of any offhand, it seemed reasonable that many of them would be scattered through the journals.....After prolonged search, with the help of many others, I succeeded in locating only three relevant studies..."²

Although there has been substantial effort put into developing better definitions and measurements of attitudes since 1963, it is, none the less, still true that surveying attitudes as an indicator of behavioral responses is at best an approximation , and in many situations misleading. This is especially true where the beliefs or expectations of other people carry importance for an individual. The whole area of quantitative methods is a highly-charged topic for many managers, and the impact of organizational climate on expressed attitudes may be substantial. Several of the statements in Exhibit 1 are charged in this way; it is hard to be confident that a respondent's expressed opinion to the statement, "the complexity of the decisions I make requires the use of quantitative methods" reflects the direction of his behavior , rather than his assessment of the expectations of his own reference group.³ The point can be overargued, but the standard questionnaire approach that is used in virtually all Factor studies has tenuous value for assessing the complex processes underlying OR/MS implementation. Even worse, a one-to-seven scale is an arbitrary calibration for measuring the degree of agreement. Throughout Factor studies, there is a meticulously exact evaluation of grossly approximate data. The correlations among the statements of Exhibit 1 have no real meaning (even with a sample size of 80) and no validity for conclusions, and yet, these are the input to a majestically subtle analytic technique. In passing, it should be added that number-crunching methods, such as factor analysis - and linear

regression, involve important assumptions about distribution and variance that can never be met in behavioral data such as opinion surveys though they are, in practice, often ignored.⁴ In general, any multivariate technique gains power at the expense of substantial requirements concerning normality and homoscedacity. It is easy to see the appeal, even the necessity, of a multivariate approach to the analysis of a behavioral system. In an exploratory program of research, where the relevant variables are hard to define, let alone measure, the data obtained is, however, generally qualitative and essentially ordinal, even where disguised in seven-point scales.

The dead horse of this Factor study has already been well-flayed in this discussion. The final point to be made is, perhaps, the most central - the others draw attention to what are really errors of mechanics, but this concerns the theoretical underpinnings of the whole approach. The Likert scales sample attitudes with no a priori reason for doing so - in essence, they are used because opinions are easy to obtain. Opinions and attitudes are manifest variables and provide a way of accessing - in principle at least - elusive and unarticulable behavior. However, implementation is a process that does not easily reveal itself in terms of structural features. Its flow may be hard to pin down, the salient characteristics may not be apparent, and its endpoints may be ambiguous (the question "what is a successful implementation?" is difficult enough, but, in many cases, simple in comparison to the one "when is an implementation implemented?") Lacking a methodology for capturing process variables, it is tempting to choose before-and-after snapshots. Lacking measurements of behavior, intentions, changes in response, etc., it is even more tempting to use attitudes as a surrogate or analogue. But by choosing to do this, the researcher may well constrain his theoretical framework; a methodology should surely support, not lead, a conceptual scheme. The choice of what to measure is an implicit statement of theory and determines the focus and attention of a study.

There are several instances in the behavioral sciences where methodological decisions have amounted to theoretical imperatives, and, in some cases, hardened into a distinct ideology. Thomas Kuhn's concept of a paradigm highlights this intimate fusion between theory and method:

"The operations and measurements that a scientist undertakes in the laboratory are not "the given" of experience, but rather "the collected with difficulty". They are not what the scientist sees - at least, not before his research is well advanced and his attention focused. Rather, they are concrete indices to the content of more elementary perceptions, and as such, they are selected for the close scrutiny of normal research only because they promise opportunity for the fruitful elaboration of an accepted paradigm. Far more clearly than the immediate experience from which they in part derive, operations and measurements are paradigm-determined.⁵"

Kuhn's "paradigm" is a theoretical-methodological template imposed on "reality" and obviously that template, in itself, governs both what data will be gathered and how it will be evaluated. The paradigm thus defines what variables are "relevant"; it is a small step then to assume that those variables that are not highlighted by the paradigm are unimportant. Some of the work of the Carnegie school in the area of organizational theory and problem-solving seem to show this latter tendency; the approach of Simon and his colleagues was initially to develop propositions that were formal enough for and phrased in terms suited to computer simulation. This position hardened in later research in this tradition to a virtual dogma of "if we can't simulate it, it doesn't exist"⁶. An even more extreme instance where a methodological choice became an ideology can be found in early behaviorist psychology. The early studies of Thorndike, Guthrie and Watson made use of rats as subjects, mainly because of the experimental control, comparability of samples and clarity of design that they provided at a time when understanding of the mechanics of elementary behavior was minimal.

By the 1940's, the behaviorist paradigm of stimulus-response had hardened to an extent that the conventional wisdom explicitly rejected all notions of "intent", "consciousness" or even "emotion" and viewed man and animal as essentially equivalent. Chomsky, in a majestic attack on B.F. Skinner, whose research represents the extreme of this position, drives home the point that Skinner generalizes from highly-constrained experiments, which by themselves, circumscribe the variables that he can examine ("Psychology and Ideology")⁷.

It is a far jump from rats to measuring attitude, but the parable holds. A Factor study may aim at uncovering relevant variables through a descriptive methodology, but the very choice of what to measure, and how, is implicitly a choice of theory, albeit at a loose level. Once the choice is made, the data of "reality" will be to a degree pre-filtered and organized within that theoretical-methodological template - obviously, this runs directly counter to the whole intent of the Factor study.

Kuhn speaks of a preparadigmatic stage that any science must pass through before it arrives at "normal" science; the latter stage involves the exploration of phenomena from the viewpoint of paradigms that have some consensus and acceptance among the scientific community as a whole (in the behavioral sciences, economics is clearly the field that has such a range of paradigms, and is thus at a stage of normal science). The preparadigm state is marked by the absence of any consensus on conceptual models. There may be several potential paradigms in competition with each other, but there is no acceptable test for comparing and validating each of them. Kuhn characterizes the research at this stage as one of more or less random fact-gathering, restricted to the wealth of data that lies readily at hand. He suggests that the social sciences as a whole, are at this stage in development (1962); whether that is so or not, his comments strongly apply to the study of OR/MS implementation - we have hardly a glimmer of consensus on a set of basic paradigms, and the most obvious general feature

of the research effort is a searching for potential schema. There is no reason to criticize this feature; Kuhn implies that it is an inevitable first step towards an applied science. One should, however, criticize research efforts that somewhat arbitrarily use a methodology borrowed from a tradition in social psychology in which these questions of definitions, conceptual schemes, instrumentation and interpretation of experimental data are far less tentative and tenacious. The Factor study is vulnerable to this charge and also ignores problems of comparability. Holt and Turner describe the limitations of political science research in terms similar to those used here; they suggest that the hypotheses tested by political scientists are generally either the loose implications of a rather amorphous theory or simply hunches about the likely outcome of some empirical study.⁸ They point out that a major difficulty for political science is that it can rarely, if ever, manipulate its variables. Because of this, and their argument is directly relevant for OR/MS implementation, the political scientist must use a comparative approach; in building descriptive schemes, he needs to identify the "background features" (Nadel)⁹ of a particular setting, isolating the general from the contingent. They emphasize that this is a key first step; in the Factor study, it is virtually ignored, and any inductive strategy of that type cannot provide the basis for comparative research.

It was suggested at the start of this paper, that examining the reasons for choosing a flawed methodology would be instructive. The aims of the Factor study are necessary and justified, especially if one compares it with the "controlled" experiment in which a specific and narrow set of hypotheses are tested. Implicitly, the Factor approach accepts the necessity for what Roger Barker terms an "ecological" approach:

"If a novice, an Englishman for example, wished to understand the environment of a first baseman in a ball game, he might set about to observe the interactions of the player with his surroundings. To do



"this with utmost precision, he might view the first baseman through field glasses, so focused that the player would be centered in the field of the glasses with just enough of the environment included to encompass all the player's contacts with the environment, all inputs and outputs; all balls caught, balls thrown, players tagged, etc.... By observing a player in this way, the novice would, in fact, fragment the game and destroy what he was seeking. So, also, he might by observations and interviews construct the player's life-space during the game: his achievements, aspirations, successes, failures and conflicts; his judgements of the speed of the ball, of the fairness of the umpire, of the errors of his teammates. But this would only substitute for the former fragmented picture of "the game" the psychological consequences of the fragments and thus remove the novice even further from the ecological environment he sought.....It would seem clear that a novice would learn more about the ecological environment of a first baseman by blotting out the player and observing the game around him.¹⁰"

The Factor study is basically ecological - it attempts to maintain a focus on the whole setting. The most important aspect of Barker's example is that the phenomenon involved is a process, and he emphasizes that a process cannot be captured through a set of snapshots. This problem is, in fact, at the core of the issue of designing a behavioral methodology.

The question underlying the study of implementation is, as was stated earlier, a simple one: "what factors influence the likelihood of success in an OR/MS effort?" The division between academic and practitioner is very blurred in this area of study (another example is Organizational Development (OD) and an opposite instance is the study of personality). The needs of both researcher and practitioner, at this stage prior to "normal" science, are parallel. The practitioner (and even more, the manager) requires better maps, for obvious reasons; the academic similarly needs maps as a basis for formulating models and hypotheses that will allow him to go beyond opportunistic fact-gathering. In many situations, the academic is also a practitioner. The best metaphor

underlying - and partly governing - behavioral studies of OR/MS should thus be one of mapping. A map is a precise schematic representation of the territory; it is also an abstraction, a simplification and a generalization. Without pushing the metaphor too far, one may suggest that this mapping process has some definite implications for the researcher's individual role as an observer and as an explorer.

There are a number of constraints impeding mapping of implementation. These have in essence already been discussed; the Factor study essentially tries to finesse them. Stated as a brief list, they constitute a formidable agenda:

- 1) Process and Contingency; implementation is a process within a dynamic system involving complex contingent variables; tracking that process is an essential feature of the Mapping study, but is substantially more difficult than examining structural, static variables.
- 2) "Success"; there is no simple definition of the dependent variable - "successful" implementation; success is often contingent on aims, expectations and situation
- 3) Data selection and instrumentation; the researcher has, at best, loose hypotheses or hunches to govern his selection of variables to examine, and there are few accepted instruments for measuring most of those variables; moreover, even the instruments that are accessible are rarely well-calibrated - in many situations the most interesting variables are the least measurable.

Each of these three headings hides a range of problems; unfortunately the development of adequate maps depends on a reasonable resolution of them all.

Process and contingency.

It is obvious that implementation is a process; it often does not have neat beginnings and endings. In some situations (especially the class of computer system for which the term "Decision Support System" is an accurate descriptor)¹¹

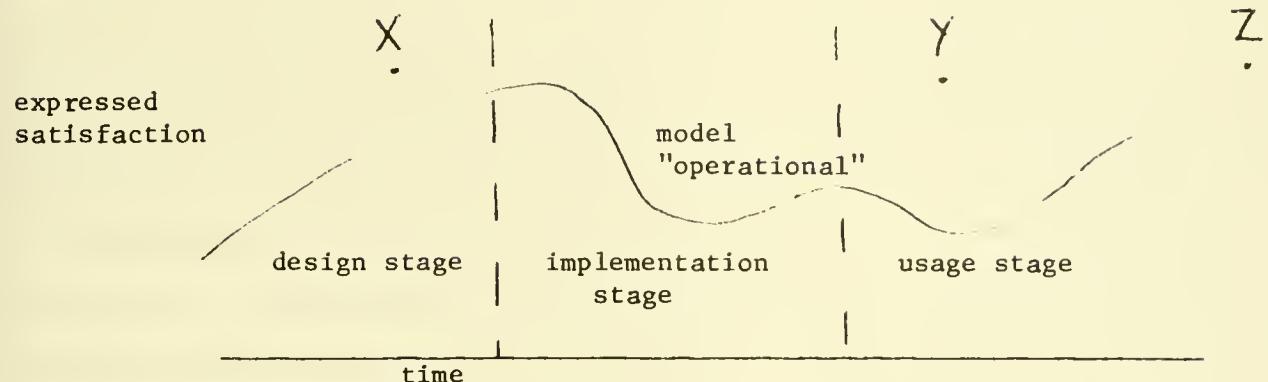
it may be very hard to determine when an OR/MS effort has been "implemented". There may be complex adaptation or evolution for some time after the system is officially released for use. Frequently, the outputs from an OR analysis in themselves generate further and different directions of exploration. The agile practitioner may deliberately include these new shifts in his approach to the manager's problem. A complex analytic project is a creation; when the group sets out, it can generally have only a tentative idea of what will be involved. Serendipity is part of even the most quantitative analyst's toolbag.

The stages of implementation, similarly, rarely split up into neat stages. Design, testing and use of the results may be highly recursive. Many researchers take this into account by making longitudinal studies. "Longitudinal" may again be a grandiloquent term for a convenient practice; such studies often involve merely taking two snapshots, "pre" and "post", and comparing the shifts in attitude or behavior between the two points in time. Such snapshots can provide insight into overt changes, but rarely do they illuminate the process. Of course, this may not matter, since many research efforts are concerned only with outputs; the locus of interest is the differences created, causally, by the OR/MS input - the process may be viewed as a Black Box. Behavioral studies, however, are almost invariably concerned with the dynamics of the intervening activity, to illuminate the Box. In that situation, much of the standard methodology of the social sciences is of little value. Cross-sectional snapshots are all too often inert dissections. More insidiously, there is a problem in timing the snapshots, a problem which is best illustrated through the specific example given below:

Managers' expressed satisfaction with an OR/MS project is often volatile and reflects expectations as much as accomplishment. The shifts in satisfaction shown in Exhibit 3 are not uncommon (the example shown assumes an analytic computer model which is to be used on a recurring basis by several managers - a

typical case would be a financial forecasting model).

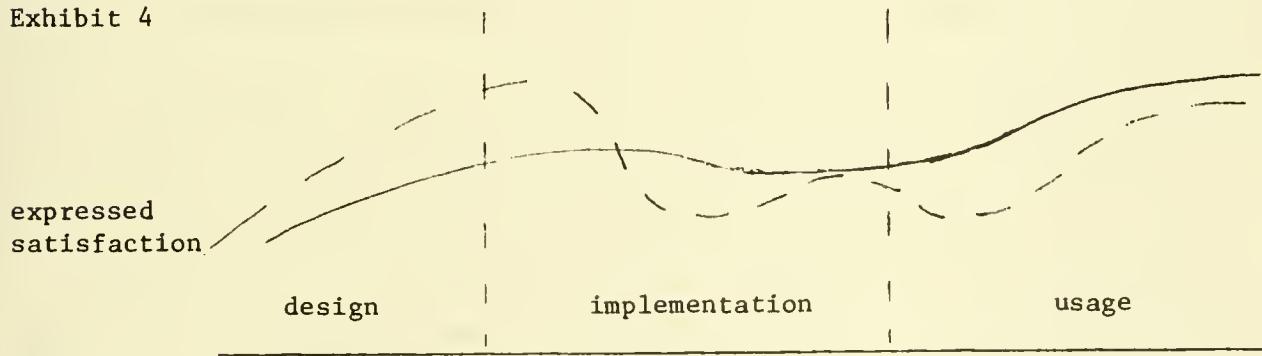
Exhibit 3



In the design stage expectations and excitement rise fairly rapidly (probably through overselling). The inevitable problems and conflicts incurred during implementation result in that satisfaction dropping and this continues after the model is made operational. There are unanticipated difficulties; managers find it hard to learn the logistics of using the model, they find that they have some needs not met by it etc., etc. If the model is a good one, it can be refined and adapted, and after these initial frustrations it becomes institutionalized. An immediate question, to raise and leave hanging, is at what point is the model "implemented" and when should its success be evaluated?¹² The most likely time at which the pre- and post- snapshots would be taken are points X and Y in the diagram (when the initial design stage has reached some definition and momentum and when the model has been in use for a month or so - four months behind schedule and when the researcher's patience and/or budget is running out). Very obviously, the results will be different if the researcher holds off until point Z. Even more importantly, expressed satisfaction - like other variants of attitude - at any single stage may reflect several other seemingly unrelated variables; in Exhibit 3, for instance, it is presumed that the project is oversold, so that ex-

pectations are raised that can not be realistically met. Lundquist¹³ defines an "oscillation" effect, where a person sees a new technique as promising immense payoffs and rushes to design possible uses for it. When he examines the technique more closely, he comes across practical limitations, costs or obstacles and his enthusiasm is dashed to a degree that he now views it with scorn, totally overreacting. A similar phenomenon is common with OR/MS; the oscillation effect is dampened down over time, but can still be volatile - in Exhibit 3, for instance, the overselling means that the user may be actively dissatisfied with the model towards the end of the implementation phase. Exhibit 4 suggests what the curve might have been had the project not raised such high initial expectations (the curve of Exhibit 3 is superimposed in dotted lines):

Exhibit 4



Even if the researcher is interested only in outputs, his use of a surrogate variable such as satisfaction will lead him into likely error if he is unaware of the impact of the users' initial expectations. If all measurements are taken at point Z in Exhibit 3, then it is true that the fluctuations - and the contingencies - have resolved themselves; there are in fact a number of excellent studies that work from a post facto viewpoint, using "hard" variables and large samples (the Northwestern group, under Radnor, Rubinstein et al have, on the whole, worked from that approach, and their conclusions are unassailable)¹⁴. However, the behavioral study that uses a single setting runs immense risks of downright

error unless the dimensions and flows of the process have been identified, even if only in broad outline.

Exhibit 3 reflects the importance of contingency; the use of an expressed attitude as essentially a dependent variable is vulnerable because the independent ones have not been correctly identified. The Factor study in effect bypasses the problem by arguing that the contingencies can be taken for granted initially and the analysis used to sift them out post facto. Such an approach can not isolate the influence of expectations - this is a "background feature" that can easily remain hidden. Gibson, in advocating a methodology of action research in studying implementation, cites a large field research project where the central, key contingency was political; he argues that the whole validity of the study depended on recognizing that fact - but it was a complete surprise to the research group.¹⁵ With no a priori conceptual base to alert the team to the relevance of this subtle political dynamic, there was no way the research group could have located its effect except through a methodology explicitly geared to describing process and identifying contingencies.

From the elevated vantage point of Mount Abstraction, it is very easy to suggest that researchers classify contingencies. To do so begs the question. The whole problem in studies of implementation is precisely that there are no taxonomies, and that the mapping process begins from a point of virtual ignorance; the main result of the research is, in fact, intended to be the sketching out of such taxonomies. This tautological dilemma is very real. None the less, it can to a substantial extent be resolved by, at the very least, ensuring that the major features of a situation are classified at some gross level. This may mean using the broad taxonomies that are available and refining them over some period of time and several research projects. This approach is one of inclusion; the gain it provides - particularly to the reader of a study - is immense.

The conclusions of even the Factor study would be much more convincing if they used this essentially comparative strategy. The drive towards classification that is common to the social sciences has, in fact, already provided the basis for such loose taxonomies. Simon's distinction between programmable and non-programmable tasks,¹⁶ and Anthony's differentiation between Strategic Planning, Management Control and Operational Control¹⁷ are two examples that have been very useful to studies of implementation. They are general and break down under too close a scrutiny, but are exactly the sort of simple, compact differentiations needed in sketching contingent features in an experimental setting. They provide reference points for the reader and for building coherence in the field as a whole. Leavitt's four clusters of variables that define an organizational system are similarly gross categories of substantial value as a quasi-paradigm; Task, Technology, People and Structure.¹⁸ They may be used merely as a checklist to describe in a few paragraphs the highlights of a research setting. The Factor study cited earlier would be transformed with even the limited additions shown below in Exhibit 5:

Exhibit 5

<u>Task</u>	60% of OR/MS projects optimization of well defined processes 30% simulation for capital expansion of complex process-flow production systems 10% financial planning models
<u>Technology</u>	multi-product corporation, capital intensive, large-scale fixed capacity process; uses basic range of raw materials; product mix and yield critical.
<u>Structure</u>	decentralized, divisional; OR/MS group centralized as corporate staff with formal charter reporting to VP for Planning
<u>People</u>	divisional managers mostly chemical and/or engineering graduates OR/MS group: 30% MBA's 70% degrees in technical subjects, math or computer science.

Exhibit 5 contains little hard data and could easily be expanded. The term "quasi-paradigm" is useful, though aesthetically monstrous. The Factor study avoids this loose organizing focus. It implies that we do not yet have any paradigms. The Leavitt schema serves in place of a paradigm; to regard it too seriously is dangerous, but it is equally foolish to assume that one can avoid any framework; the quotations from Kuhn and Barker make that clear. In the tradition of Germanic scholarship, one could glorify it with the description "comparative relativism". It is not in any way a theoretical model, but a focus that elevates the explicitly atheoretical Factor study to the type of comparative research Holt and Turner discuss in relation to political science. Given the aims of the mapping process, it seems a necessary first step that bounds the data and its interpretation within a contingent and therefore, relativistic frame. Holt and Richardson's comments on the development of paradigms in comparative politics seem directly applicable:

"We find it difficult to see how the application of an analytic routine, however elegant, to data selected on the basis of explicitly nontheoretical considerations can contribute significantly to its accomplishment" (i.e. to the emergence of paradigms).¹⁹

One final point of modification needs to be made. The role suggested here for contingent quasi-paradigms is at a relatively modest level; it provides for a set of criteria for analysis and evaluation that is less formal than explicit hypotheses, but that amounts to a template to impose on an undefined reality. It is in no way argued here that comprehensive taxonomies can ever be developed. Each of Leavitt's four tidy categories branches out virtually to infinity; one could never even adequately classify Task, the most definable of the four. The People category, in particular, has defied classification, and years of psychometric effort has generated tests whose reliability and meaning is consensual rather than proven; those tests also use a series of molecular labels, such as "extrovert" or "intelligence quotient", that are as much challenged as used.²⁰ Studies of implementation are un-

likely to do better than the rest of the behavioral sciences, whose definitions are qualitative and always to some degree arbitrary. Occam's razor must be the researcher's guiding principle; in developing taxonomies the underlying question is "what are the fewest variables at the broadest level of detail that must be defined if one is to adequately capture the contingencies of a situation?" The recommendation here is part of an outline for a methodology, the exact application of which must obviously start from a theoretical base, meshing with the specific aims of the research effort and requiring ingenuity, commonsense, trial and error.

"Success"

The whole field of implementation research will take an immediate step forward if a comprehensive taxonomy of OR/MS "success" can emerge. Success is the key variable in all the research and, yet, it is extremely difficult in most cases to either define or measure it. In a few situations, success can be identified in solid terms - an OR model "works", is used and management feels that it led to better decisions and eagerly asks for more. Huysmans cites three "levels of adoption" for OR and indicates that these may be in conflict or overlap:

1. "Does the development of an OR model lead to management action?" (Is the model used at all?)
2. "Does the development of an OR model lead to management change?" (Is full value from the OR model received?)
3. "Does management's contact with the OR method lead to recurring use of the OR approach?" (Has the approach become an integral part of the management thinking?)²¹

He cites some rich case examples, where it is very difficult indeed to determine either the degree of success, or even of implementation. The researcher has to choose some anchor point and generally this will be a surrogate for success, such as management satisfaction, the percentage of projects completed, or the number of hours of

terminal usage. There are some contexts where even a surrogate may be hard to find. For example, a model or analysis may be only part of the input to a management decision; this is especially likely to be true for major strategic decisions where the manager involved will, in the last resort, integrate in his own mind data and recommendations from a variety of sources, of which the Monte Carlo simulation is only one (Little's paper "Models and Managers; a Decision Calculus" stresses this point²²). One can rarely isolate the impact of the OR/MS project on that decision, nor then assess the value of that venture in improving the quality of the decision. More subtly, there are some situations where an analyst may best serve his client by not implementing any quantitative approach, but by helping him clarify the problem. At the other extreme, he may "successfully" implement a technique that does not lead to a better decision; this is true for many portfolio selection models for R&D projects, where the manager can outperform the model.²³

The most important scenario , where formal measures of success may be totally misleading, is also one of the most productive, and is a not uncommon experience for competent analysts. Here a manager requests assistance with only a very rough idea of what he expects. He may define the problem and request some formal model. Later, he recognizes that his assessment of the problem was in some way incomplete or inadequate. His interaction with the analyst may lead him to examine his own perceptions or assumptions, extend his understanding of his environment and point to perspectives that he was unaware of. By the time the process is complete, he may view the model as a poor one and scrap it. The OR/MS effort thus appears a failure, but it has contributed to his learning to an immense degree. By any criterion, the project is a success, but there is no easy measure that identifies it as such (although Huysmans' third question includes it: "Does management's contact with the OR method lead to recurring use of the OR approach?"; the problem here is that this type of success is apparent only from a long time perspective, and that the success is difficult to associate with a specific project). This dilemma again highlights the

fact that much of implementation is a process; process variables must be tracked over a meaningful time-frame and can rarely be captured through structural analysis or from examining outputs (such as the actual use of the model). There are some instructive parallels here with the study of organizational change programs; these will be discussed later, and indeed form the base for some of this paper's specific recommendations.

Obviously, "success" is yet another contingent variable. It depends partly on intent and aspiration. It may also reflect prior expectations or management capabilities (or their absence) rather than the OR/MS specialist's actions. The whole drift of this paper is to suggest that any formal measure of some "hard" variable, such as terminal usage, will be inadequate in most cases, even though convenience often leads to the use of such a surrogate. Huysman's three questions are useful as a framework for a contingent assessment. This author finds it extremely helpful to distinguish between Service and Product; the differentiation captures both the intent and the process of a specific implementation and for that reason is worth some discussion here.²⁴

Frequently, a manager has a clear idea of what he wants from an OR/MS project; it may well be that he intends to act on the basis of the analyst's recommendation (e.g. in a risk analysis of a capital investment program). Here the criteria for success are relatively easy to identify; the manager has expectations and aims for the project and these will either be met or not, so that his perceptions - and expressed satisfaction - are an acceptable measure. Moreover, the effort can generally be evaluated soon after its implementation - and that implementation point is similarly well-articulated. It seems accurate to define this effort as one of Product. By contrast, there are many instances where the analyst provides Service. The manager draws on an expertise that he feels is of relevance for him, but does not look for a specific output; he may even, in fact, use this expertise on an ^{or} opportunistic basis. The impact of the OR/MS effort on his actual decision may thus be tenuous

with no product being generated. Frequently, an analyst deliberately cultivates a manager in this way, building credibility and "trust".

This distinction between Product and Service is critical in several ways. It may partly account for Churchman's important finding,²⁵ confirmed by innumerable other writers, that many "implemented" models are never used by the manager for whom they were designed, even though they are technically sound and well-fitted to the problem involved. Churchman set a student group to review six years of issues of "Operations Research Quarterly" and concluded that there was no evidence that any of the models discussed were ever used. He points out that without a deeper understanding of the manager's true needs, the analyst is likely to provide, in the terms of this paper, some Product designed from his own largely normative assessment of what is required by the problem. Obviously, there are many factors involved in the phenomenon Churchman describes, but in at least some instances, it is caused by the analyst providing Product where the manager implicitly requires Service. Where that is so, the venture is obviously not a success, however elegant the model may be, and however impressive the journal article it thus provides. In a Service project, the absence of a tangible output is not necessarily a failure. The success of such an effort can be measured only in terms of process variables; there are several such that may be suited to specific situations, such as the manager's improved sense of comfort with the OR/MS approach, or the strengthening of his relationship with the OR analyst. Since the manager may lack the perspective to articulate these feelings, even to himself, at the time the project is complete, they may not be directly measurable, except in terms of his future willingness to use OR/MS methods. The central issue for the researcher is once again one of contingency; without an understanding of the full Field (in Kurt Lewin's sense of the term²⁶) he cannot choose a measurement or a time-frame for determining the degree of success.

Data selection and instrumentation

The straw man example of the Factor study discussed the problem of collecting behavioral data, and the danger of surrogate measures. It is worth reiterating , briefly, the main argument there. What to measure is the first choice the researcher has to make. The second decision is dependent on this one; how to do so. The tendency to reverse these two procedures severely constrains the conceptual focus of any study. It also predetermines the type of conclusions that can be drawn from the data. The use of a particular sampling technique and of ratio scales implicitly assumes a set of rules (mainly statistical) as to what constitutes admissible evidence for evaluating that data. Where these rules are broken, the study can be criticized and probably rejected on that point alone. This author suggests that qualitative behavioral data is rarely suited to the rules of multivariate analysis such as linear regression, factor analysis, or multi-dimensional scaling; to an extent this statement reflects a personal bias that can be challenged, but the reader can himself decide if such data meets the following rules, all ofwhich are critical for the techniques mentioned:

1. scores are normally distributed and not skewed
2. the data is at least on an interval scale
3. the sample size is substantial
4. the sample is random

The example of an attitude survey used earlier will serve once more. On a 1-to-5 scale, where a respondent circles a "1" for "strongly disagree" and "5" for "strongly agree", the aggregated results are likely to be skewed, the scale is surely ordinal in nature, and it is rare for the sample size, in a detailed implementation study of a single or several settings, to reach the range of about 120 needed for factor analysis, especially, and it is clearly in no ways random.²⁷ Ignorance and convenience explain the use of multivariate techniques where they are not justified, but far more responsible is the fact that a rhetorical device has hardened

into a convention, just as Greek oratory ossified so that techniques like litotes became expected in every speech. Statistical conventions easily become fixed. Given the lack of definition and precision inherent in most behavioral research, the use of elegant flourishes like multi-dimensional scaling does not support the rhetoric of a study, but is an artificial intrusion. Statistical method is obviously essential in behavioral research, but when the technique is cautiously selected to fit the demands of the data collection process it is much more likely to be nonparametric and distribution-free. The author feels strongly that the researcher in this context should venture into the preparadigmatic darkness shielded from temptation and peril by Sidney Siegel's superlative "Nonparametric Statistics for the Social Sciences".²⁸

Towards a Behavioral Methodology

The preceding arguments have, on the whole, pointed to difficulties and inadequacies, with hints as to a desirable methodology. Clearly, the specification for such a methodology must leave ample room for the exact demands of a particular study, and also for the exact demands of the researcher himself. The list shown in Exhibit 6 below, presents the essential features of the approach to be recommended here - the list is a bald series of descriptors, each of which is expanded and defended in the rest of this paper.

Exhibit 6

<u>1. The role of the researcher</u>	diagnostic involved observer opportunist
<u>2. Field description</u>	historical comparative contingent
<u>3. Measurement</u>	referential validation via perspectives rhetorical

Exhibit 6 (continued)

4. Interpretation and conclusion

"mid-range" theory
mapping
generative explanation

An implicit assumption, that hopefully does not require further justification here, is that the behavioral study of implementation is essentially one of field research and that the researcher is an observer who is in very direct contact with the elements and individuals concerned during much of the implementation process itself. This factor has some important implications for the researcher's role, which have been described in other contexts (especially in sociological research)²⁹. To a large extent, a study of implementation is forced to be in the field research tradition; the experimental approach requires some degree of control over the choice and types of settings to be used, as well as the ability (conceptual and/or practical) to focus on a few, well-defined variables. In studying implementation the researcher rarely has such control. He has limited access to rich settings and is generally opportunistic in selecting sites for research. In many instances, he depends on a consultant or academic who is strongly involved in the implementation to invite him in - sometimes he, himself, recognizes that an implementation effort with which he has some connection offers a chance for a behavioral study at the same time.

Because of these factors, the researcher is usually an involved observer, known by the members of the organization to have some stake in the implementation itself, even if only as one of the "guys from Transylvania Tech." The whole question of participant observation has been discussed in many books on field research; some writers have stressed that the observer himself is part of the field - his presence is a disturbance, even if only a small one.³⁰ One tradition turns this from a weakness to a deliberate strategy of Action Research in which the observer enters wholeheartedly into the field giving help ^{and} _^ advice, responding to his own biases and perceptions as the only way to get an existential awareness of the processes he is examining.

Although the cool, dispassionate researcher on OR/MS may disavow that approach on principles unscientific, there are many instances where he is close to that position; he may, for example, be involved in the design of the project and thus take some lead in defending it in times of difficulty and in smoothing out problems encountered during implementation. There are some obvious limitations to Action Research, especially with respect to drawing useful general conclusions from the experience. However, the researcher should not pretend to an objectivity and absence he does not have; in particular, in field research, where he is dependent on the goodwill and readiness of the members of the organization to share their perceptions, any effort to stay uninvolved may damage his study. The following extracts from a discussion of sociological research summarize issues that must be consciously resolved by the individual researcher:

"the presence of a stranger, particularly an observer, in a natural human situation introduces some measure of disturbance in the scene....Any method that might be used to study situations like these poses special problems; if field research creates disturbance, other methods tend to create artificiality.

Observing without being observed is virtually impossible to manage in natural social settings. The need to sit in on relatively private discussions, and to ask questions, precludes this tactic as a reasonable option..... the spectre of a relatively impassive observer, whether or not taking notes, barely showing appropriate effect or active curiosity, and offering few if any cues as to what he is "really up to", can be very disturbing to the hosts.

The data that are obtained through the observer's judicious, but active intervention might not actually be readily available to perception and conception were he not actively there. For example, the careers and ideologies of the hosts are not demonstrably there without his being there; they are created or revealed in and through the interaction of the hosts with the observer.^{31"}

These issues have always been central to the study and implementation of planned change. OR/MS methods are frequently described as "change agents" and the role of the analyst is not dissimilar to the role of the consultant in OD (Organizational

Development) defined by Schein, Beckhard, and others.³² It is worth noting that some of the most interesting - and convincing - articles recently on the topic of OR/MS implementation have explicitly analyzed the process in terms of social change.³³ Without suggesting that OD and OR/MS implementation are directly analogous, the research and practice of planned change provide valuable conceptual frames and guidelines for the researcher's own role. (Dimensions not well discussed, however, in that literature are Task and technical or technological change; thus any comparison between OD and OR/MS has reciprocal value for both those fields of study).

Schein, building on Lewin's early model, identifies three distinct phases in any change process:

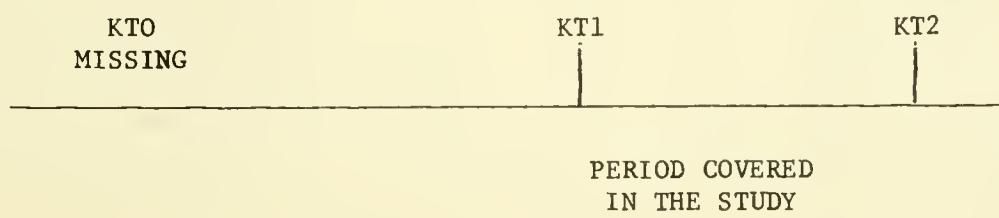
- "1) Unfreezing: an alteration of the forces acting on the individual, such that his stable equilibrium is disturbed sufficiently to motivate him and make him ready to change; this can be accomplished either by increasing the pressure to change or by reducing some of the threats or resistances to change.
- 2) Changing: the presentation of a direction of change and the actual process of learning new attitudes.....
- 3) Refreezing: the integration of the changed attitudes into the rest of the personality and/or into ongoing significant emotional relationships.^{34"}

Schein's description is clearly applicable to the change processes implied in OR/MS implementation, though that is only incidental to the argument here. (Sorensen and Zand's paper which specifically applies the Lewin-Schein to OR/MS is strongly recommended both for its research design and its conclusions³⁵). Schein defines the skills required by the "process consultant" whose role is to facilitate change in an organization; he stresses the importance of knowing "how to diagnose and establishing helping relationships"³⁶. The consultant must have a theoretical and practical knowledge in the six main behavioral processes Schein identifies as affecting organizational performance:

- 1) communication

- 2) member roles and functions in groups
- 3) group problem-solving and decision-making
- 4) group norms and growth
- 5) leadership and authority
- 6) intergroup cooperation and competition.

This model of social change has value in itself for studying implementation, but much more important is its implications for the role of the researcher. The field observer in an OR/MS study is not necessarily a consultant, but he is likely to make - or be expected by others to make - a contribution towards facilitating implementation. He is thus involved in the field from the onset. The skills of process consultancy described by Schein neatly mesh the functions of observer and facilitator and amount to an approach to research that is scientific and at the same time that accepts the impossibility of detached "objectivity". The most important of these skills is the ability to diagnose. The diagnosis is one of process, not of structure. It begins with getting some "feel" for the situation and is basically historical in its approach, in that the observer aims at explaining and tracking change. The historical focus is in no way equivalent to the longitudinal study; the difference between the two is shown by Forcese and Richer, who cite as example a conventional experimental design using two groups, "control" and "treatment". In the longitudinal study, two measurements are taken at KT1 and KT2 ("Knowledge Time" 1 and 2) and the difference between the points is ascribed to the causal variable controlled for in the two settings. However, there is a missing Knowledge Time, KTO, that is historical:



It may well be that the causal variable is obscured within KTO. Forcese and Richer, in comparing performance of pupils in schools which have "resource" teachers, suggest that:

"if an individual entering a school with resource teachers is there because of some prior commitment to learning, this in itself may cause a greater positive change in his performance, although its effect is likely to decrease in time. The longitudinal study does not completely take care of this possible inequality in motivation or commitment.³⁷"

(Exhibit 3, earlier, makes a similar point in discussing the timing of "snapshots").

The past, KTO, is often not directly accessible. In the study of school performance, it may be feasible to reconstruct it from records accumulated routinely by the organization; Webb et al strongly recommend using these "unobtrusive measures"³⁸ However, the diagnostic approach can at least approximate KTO even where records are unavailable or uninformative. The researcher's whole frame of reference, using that style of observation, is to define the main forces acting on the field that is of interest to him. He cannot help but ask what are the key features of the situation, how did they arise and what factors most contributed to them. In diagnosing OR/MS projects, this approach requires identifying not just what management's current attitudes are, but also the dynamics underlying them. This argument is, of course, somewhat idealized, but there is a qualitative difference between the neutral, experimental stance assumed in most studies in the social sciences and the necessarily historical frame of diagnosis. The first and in some ways central recommendation for a behavioral methodology for OR/MS studies is simply to accept the research style implied in Schein's description of process consultancy; it is liberating, more convincing as a role definition and it is based on some very real successes in studying elusive processes of behavior change.³⁹ Every science leans heavily on idealized myths; in the "hard" sciences, these myths have generally been ones centered around the lonely, very precise analytic researcher alerted to the implications of tiny clues in his data. The organizing myths in OD are those of a warm, empathic guru

helping others arrive at the unexpressible. The cold realities of task and impersonal analysis involved in OR/MS make this latter archetype unconvincing for the type of study discussed here, but it seems very arguable that Schein has more to offer as a research guide than all the Factor Students combined. The pseudo-physicist model of research that underlies (probably unconsciously) many studies of OR/MS is not an adequate way of looking at a rich and complex reality.

The diagnostic approach is opportunistic. Gibson's experience, cited earlier, is not untypical. He had only broad expectations as to what he would find in his field setting. Of course, he was also sensitized to certain types of phenomena; the diagnostic approach is biased inevitably by the quasi-paradigms governing what data is initially collected. None the less, the diagnostician (the example of a clinical psychologist is valuable here) is continuously looking for clues. He assesses the field systematically, but has to be attentive to discrepancies, contradictions and new implications in the data he gathers, and to use them as opportunities. Wherever possible, this opportunism should be rooted in measurement rather than personal hunch. The field description that is the result of his diagnosis essentially clarifies and structures the main forces in the field. It draws on taxonomies where available and suitable, both to understand and to classify.

Obviously, a full description of the field is impossible in any literal sense. The whole process of historical research of any sort largely involves a continuous redefinition of fields that may have been described literally hundreds of times. For example, any essay on the causes of World War II includes a field description which is diagnostic, historical and comparative - the very features suggested here as desirable - but A.J.P.Taylor's description differs immensely from H. Trevor-Roper's and his in turn is different from J.C.F.Fuller's.⁴⁰ The "facts" are well-known, but assembling them into a coherent whole, of necessity, requires interpretation and filtering. Hammond⁴¹ points out that it is very easy indeed to accept the need for a taxonomic, contingent description of the whole situation in examining OR/MS and even

to sketch out the main headings for the variables to be included; however, the next step, expanding these headings and articulating the relationships among the many variables, is a bewildering and self-defeating exercise. His point is a central one; the whole argument presented here is valueless unless it includes suggestions as to how this vital step can be made. To a large extent, the diagnostic approach bypasses some of the problems Hammond describes, although at a cost; it attempts to define, mostly in qualitative terms, the key dynamics of a situation and relies on the researcher's skills in drawing on a range of theoretical models as seems most suitable to him, rather than on a priori conceptual schemes. Of course, this approach has to be made operational. Gibson's study provides a specific example of how that can be done. He and a team of students spent a full six months on what he calls "initial data-gathering", but which more closely corresponds to field diagnosis:

"First, we documented the decision-making processes in branch site selection and branch performance appraisal. Second, we accumulated qualitative information on characteristics of key individual officers, including such things as apparent personality traits, cognitive and decision-making style, attitudes towards the banking business and bank strategy, and attitudes towards the use of computers and the particulars forthcoming simulation model itself. Third, through the interviews, file data and some previous research in the bank by others we developed a fairly comprehensive history of the careers of individuals and the evolution of the social system of the organization. Finally, we learned about the structure and formal organization procedures of control, rewards and planning.⁴²"

The most important aspect of this initial data-gathering is its reliance on qualitative observation (particularly interviews) focused on historical factors. In this way the team developed a knowledge of the processes leading to the present structure or equilibrium, and not just of the structure itself. Essentially, this constituted the Knowledge Time KTO missing in the longitudinal study. It was also opportunistic in that the focus on trying to make sense of, rather than just measure features of,

the situation, alerted the team to several apparent contradictions between the formal processes of the organization and actual practice. Gibson was able to pick these up (and use them as the main impetus towards the final conclusions of the study concerning political forces affecting implementation) mainly because he had not constrained the possible range and type of data through predetermined "controlled" instruments. In a particular instance, the use of an a priori conceptual focus would have obtained confirming and invalid results:

"We learned in considerable detail from Tom, a staff officer, how he went about searching for potential branch sites, gathering information on alternative sites, investigating the possibilities of buying small banks in promising areas versus opening a branch de novo....We were able to flow chart the decision process just as he had laid it out. It appeared to fit a rational decision-making model quite well. Then we talked to the executive vice-president. Routinely, we asked "how does the bank go about opening branches ?" Excerpts from his reply follow:

"Well, we pretty much open them wherever and whenever we can....I heard the other day, the president of the little Bridgeville bank had died, I got on the phone and called his son, offering my condolences and saying I hoped he wouldn't feel I was overstepping the bounds of decency if I asked if he had considered putting his bank up for sale..."

Our next question was what Tom, down in marketing, does. The EVP's answer was "Well, he checks out the prospects, works out the figures, satisfies the regulatory requirements for us." In other words, what had appeared to be a straightforward decision process of intelligence, design, choice by a middle-level staff man, in fact, turned out to be a somewhat opportunistic, almost random initial process.^{43"}

There are several object lessons to be drawn from this illustration, not the least important of which is that looking at any aspect of an organizational process from a single perspective is likely to be dangerous. Gibson's lucky find is a commonplace of good field research. The organization itself is often unable to articulate

its own activities - both Tom and the EVP see only part of the phenomenon and from a very personalized viewpoint. The resolution of such differences and partial explanations can only come from the researcher's own effort to fit the various symptoms together.

Gibson describes his approach in terms of a case study. It seems more accurate to define it as "case-centered". The assembling of anecdotes, opinions, and attitudes that generally constitutes a case study is of trivial value in research. It raises immediately the question "so what?". There is a second step required to evaluate the diagnosis stage, generally through selecting a narrower focus and the testing of explicit hypotheses. The preliminary step is essential. It provides the reference point for the second one, in terms of a basis ^{for} _A refining hypotheses, defining suitable instruments for data collection and justifying the conclusions to be drawn from that data. Essentially, Hammond is correct in saying that the field cannot be adequately described. What can be done is the broad dynamics identified and explained - and more importantly - justified to the reader. This initial step then permits either a detailed examination of some single variables within the field or the exploration of causal links among variables. The mapping analogy used earlier is a useful one. No general map can be developed, but outlines can be sketched and experiments then made which either fill out some of the details or which test to see if the outline does seem to hold when viewed closely from other perspectives.

The second stage is clearly separate from the initial field description and is avowedly scientific in a way that a case study can never be. In most cases, this "evaluation" step involves formal testing of hypotheses. It might be argued that where the researcher has a clear conceptual model to be tested (as in Sorensen and Zand's use of the Lewin-Schein scheme discussed earlier)⁴⁴ the initial field description is not necessary, and that such a research study is in the area of "normal science" defined by Kuhn.⁴⁵ The key argument of this paper is that it is necessary even in those cases (which are unusual in OR/MS implementation studies anyway). In the behav-

ioral study, it is rarely apparent that the conceptual model used is applicable to the contingencies of the setting. An extreme instance of the lack of applicability might be a study which examines the relation between organization structure and implementation success using the hypothesis that the likelihood of implementation is enhanced where the OR/MS group has a formal charter and is centralized as a corporate staff function. Regardless of the richness of data and analysis generated in such a study, there is a need for a touchstone by which to assess the choice of scheme. If, for example, the preliminary field description showed a historical pattern in any of the research sites of substantial shifts in technology, new uncertainties in markets and important changes in environment, then the study would need to justify - and in this instance seems unlikely to be able to do so - the choice of organization structure as the central focus. There is no evidence that structure is important in the dynamics of this field. If in fact, the researcher is able to show that his chosen approach suits the field, the initial diagnosis has value in providing for comparative conclusions. His hypotheses can be evaluated in relation to the contingencies and one may then identify the extent and circumstances under which the conclusions apply. Generalizing from a particular study to a more global or general set of rules can only be done in this way. Omitting either the field description or the evaluation stage limits the potential value of any research effort of the type considered here. Omitting the second step reduces it to a case study, which can be at best provocative and more usually, anecdotal (this is not to discount its usefulness for teaching purposes; in addition, one indication of the tentative nature of OR/MS studies is the extent to which one welcomes good "war stories"). The more common practice is to omit the first stage; the main focus of this paper is to argue against that tradition.

The referential nature of the field description is also of central value in choosing a set of instruments for collecting data, regardless of the specific hypotheses to be tested. It allows the researcher - and later, the reader - to evaluate

what instruments are to be used. The field description should fulfill that function. For example, the choice of an attitude survey in a particular study is clearly justified if the field description shows that the organization, department, or group of interest contains individuals who differ in definable ways and whose differences appear to have had substantial impact on the organization's strategy or activity. Essentially, the diagnosis states that "this is what matters in this situation". The attitude survey is then rhetorical in the sense that it allows one to examine implementation in terms that are relevant to "what matters". If, on the other hand, the field description identified the main contingencies as being concerned with environment and uncertainty, then obviously the attitude survey is rhetorically unsound. This criterion for assessing the use of instruments for collecting and for evaluating data cannot be overstressed. There are many behavioral studies in all areas of the social sciences where there is no evidence that the researcher has even considered whether or not linear regression is suited to either the nature of the variables collected or the general argument underlying the work. The reader may be skeptical, but it is only the provision of a field description that can permit him to resolve his doubts.

Karl Deutsch defines Truth as "a relationship between different streams of evidence. A statement is more likely to be true, the larger the number of different classes of evidence that confirm it.⁴⁶" Deutsch's epigram provides the last main recommendation for a behavioral methodology. Even where a measuring instrument is rhetorical, it is still fallible. By its very nature, a single instrument selects a single perspective on an often highly systemic phenomenon. Barker's discussion of an Englishman studying baseball is directly applicable here. He points out that "measuring" the first baseman, no matter in what depth, will not provide an understanding of The Game. However, measuring the first baseman and then measuring the batter allows one to compare and contrast and, by that very process, to resolve the details into a larger whole. In addition, the focus on the batter is valuable in

checking out the measurements of the first baseman; it may merely confirm those measurements, which means it provides a second "class" of evidence, an additional perspective. Behavioral measures are at best approximate, and if they are surrogates, such as attitude surveys, they run a distinct risk of downright error. Here again, the dilemma is caused by the fact that behavior is a process that rarely conveniently stands still. Any snapshot of a moving phenomenon is unidimensional; just to take another snapshot from a 90-degree angle adds an extra dimension. This author used that approach in a study of cognitive style, only because of obvious loopholes and limitations in the available methodology. In retrospect, the strategy seems an excellent one even where one works from a position of strength and confidence. The study aimed at developing a classification of individual differences in thinking style. A series of standard pencil-and-paper tests were given to a sample of over 100 graduate students. The results were analyzed in detail and seemed to support the hypotheses that had been formulated prior to the study. However, it was clear that the support was decidedly circumstantial; the tests pointed to differences among subjects that were consistent, but which could, perhaps, be ascribed to academic skill, verbal versus visual acuity, etc.. A set of subsequent experiments were designed that used the classifications developed. One of the aims in each experiment was to eliminate other plausible explanations of the initial results and to generate further circumstantial evidence on the basis that if the classifications were valid, then predefined groups of subjects should differ in X respects. In this way, it was possible to indirectly control for (or at least account for) such variables as motivation, academic background, skill versus habit, personality, etc. The important point is that no one of the experiments was conclusive (although they all had the .05 or .01 level of credibility provided by statistical inference); the use of a range of contrasting perspectives on the same phenomenon built up a more convincing whole picture and helped to answer the many questions raised by any one of the experiments. In general, no one type of measurement of any variable such as motivation or

attitude can be regarded as sound. An additional measurement of the same or related variables is essential as a check and valuable as giving depth to an otherwise flat picture.

The aim of any experimental method is to permit the development of useful conclusions. The behavioral study of OR/MS implementation is in the tradition that the sociologist Robert Merton has called "theories of the middle range".⁴⁷ LaPalombra summarizes what the term implies:

"I take Merton to mean that empirical research in the social sciences should avoid a theoretical fishing expedition and pretentious impossible attempts to "test", say, the propositions generated out of Parson's four-sector description of society and its subsystems. More specifically, Merton seems to be saying that comparative research is likely to be trivial unless the propositions one is probing empirically give us some (perhaps intuitive) reason to suppose that our findings will make the creation of general theories less impressionistic or deductive than they now so obviously are.⁴⁸"

Essentially, this calls for a mapping approach to research. It also requires generative explanation. Here again, in studying implementation, the field description is the touchstone. The researcher at some step in the argument is solely concerned with showing that his data supports his hypotheses, using the often arbitrary but independent tests of statistical significance. That step has to be completed, but it is not in any way an endpoint. It permits conclusions to be drawn. Those conclusions are of two types, both of which may be provided by a single study; the first type is of the form "under such and such contingencies, X applies", and the second "in general, X applies". The case study is limited to the first contingent type of conclusion. Without the meticulous diagnosis of the field recommended earlier, many studies derive general conclusions where only the first type are really valid; this is, once again, the distinctive weakness of the Factor study. In some instances, a Factor researcher may implicitly acknowledge this and hedge; he will end his paper with some statement that the research is preliminary and points to some interesting phenomena that will

require further studies taking into account.....This hedging is in no way a "generative" explanation; it adds nothing to our ability to make maps and is better left unstated - too many trees have died in vain littering journals with such inconsequential research. Both contingent and general conclusions are of value; the first is obviously more limited, but if a study points to a meaningful phenomenon and spells out the contingent factors that can be so far linked with it, then it does provide insight and impetus for further research. The field description is the essential base on which to set conclusions, both contingent and general. The explanations needed in behavioral studies are empirical and of the middle range. The recommendations of this paper attempt to synthesize and compromise between the two extremes of the atheoretical case study and the arealsim of the experimental tradition. The recommendations are relativistic and skeptical; we lack certainty and the burden of proof is on the researcher in a way that is not true in the "hard" sciences, where accepted paradigms and assumptions permit one to take a range of important considerations for granted. The aim of any single behavioral study msut be generative - to contribute to the wider field of research. The arguments made in this paper are not intended as absolutes; the key point to be made is that methodology must support the aims of research, not lead it, and any single effort should concentrate on designing a method that is convincing and generative. Within these two guidelines is ample room for ingenuity and exploration.

Conclusion

If the preceding arguments are not convincing, no additional verbal flourish here will make them so. In concluding the paper, it seems best to summarize the operational steps involved in the methodological approach recommended here. The author suggests that this summary is not only a checklist for the researcher, but also for the reader of any research study.

Exhibit 8

Step 0: Develop a viewpoint (i.e. do not assume the Factor study pose)

Step 1: Diagnosis and Field Description

- a) Entry: immerse self in field, develop credibility as an involved observer.
- b) History: identify key contingencies and historical events, patterns contributing to them.
- c) Field definition: ensure that the field is fully defined, and that Task, Technology, People, and Structure are all sketched in at the broadest level of detail that seems relevant.
- d) Field classification: categorize the main features of the field, using taxonomies, to permit comparative conclusions and generalization

Step 2 Evaluation

- a) Select focus: Refine hypotheses and/or conceptual focus in relation to the field description
- b) Select data collection instruments: ensure data collected takes into account the contingencies and dynamics previously identified
- c) Select additional perspectives: determine the weak points, potential questions raised by, the d.c.i. in (b) above, develop reinforcing, contrasting perspectives
- d) Select statistical methods: from the field description and the data collection instruments chosen, determine what statistical method is suitable and convincing; demonstrate that critical assumptions involved are met (eg. homoscedacity, normality, etc.) and that the precision of measurement and reliability is adequate for the chosen method.
- e) Analyze data: evaluate data and hypotheses

Step 3: Conclusion

- a) Assess evidence: from the statistical analysis, draw those conclusions that are supported statistically and show that they are convincing and useful in relation to the field description.
- b) Identify implications: relate the conclusions to the contingent aspects

(continued)

- b) of the field; draw comparative conclusions; identify what aspects of these conclusions can be generalized across all settings

Step 4: Communicate: if the research report ends by saying that nothing is proven, but that the study points towards interesting lines for future research, don't publish.

REFERENCES

1. Armstrong, J. Scott, "Derivation of Theory by Means of Factor Analysis or Tom Swift and His Electric Factor Analysis Machine", *The American Statistician*, vol. 21, No. 5 (December 1967) pp. 17-21.
2. Festinger, L., "Behavioral Support for Opinion Change", *Public Opinion Quarterly*, 28, pp. 404-417. 1964.
3. Ajzen, I. and Fishbein, M., "Attitudinal and Normative Variables as Predictors of Specific Behaviors", *Journal of Personality and Social Psychology*, vol. 27, No. 1, pp. 41-57, 1973.
4. See Blalock, H.M., Social Statistics, McGraw-Hill, 1960, for an excellent summary of these and related issues.
5. Kuhn, T.S., The Structure of Scientific Revolutions, University of Chicago, 1962, p. 125.
6. See Argyris, C.A., "Some Limits of Rational Man Organization Theory", *Public Administration*, May 1973. Argyris' critique focuses on the ethical and social implications of the Carnegie School's research, but its points are very compatible with the argument made here.
7. Chomsky, N., For Reasons of State, Vintage Books, 1973, Chapter 7.
8. Holt, R.T., and Turner, J.E., The Methodology of Comparative Research, Free Press, 1970, p. 70.
9. Nadel, S.F., Foundations of Social Anthropology, London, Cohen and West, 1951, p. 229.
10. Barker, R.G., Ecological Psychology, Stanford University, 1968, Chapter 1.
11. See Scott-Morton, M.S., Management Decision Systems, Division of Research

11. Harvard Graduate School of Business Administration, 1971 and Scott-Morton, M.S., "Decision Support-Systems", The Design Process, Sloan School of Management working paper # 686-73, 1973.
12. Huysmans, JHBM, provides an excellent discussion of the meaning of 'implementation' in "Operations Research Implementation and the Practice of Management". unpublished paper, Nov. 1973.
13. Lundquist, M., Sloan School MIT personal communication.
14. Neal, R.D., and Radner, M., "Relation between formal Procedures for Pursuing OR/MS Activities and OR/MS Group Success Operations Research", Vol. 21, No.2 March 1973, pp. 451-474. and Rubenstein, A.H., et al, "Some Organizational Factors Related to the Effectiveness of Management Science Groups in Industry", 13, 1967, pp. B508-B518.
15. Gibson, C.F., "A Methodology for Implementation Research", unpublished paper Harvard University, 1973.
16. Simon, H.A., New Science of Management Decision, Harper and Row, 1960.
17. Anthony, R.N., Planning and Control Systems; A Framework for Analysis, Division of Research, Graduate School of Business Administration, Harvard University, 1965.
18. Leavitt, H.J., in New Perspectives in Organizational Research, edited by W.W. Cooper, H.J.Leavitt, and M.W. Shelly, Wiley, 1964, pp. 55-57.
19. Holt, R.T. and Richardson, J.M., Competing Paradigms in Comparative Politics in Holt and Turner, op. cit.(8), pp. 67-68.
20. Whatever consensus the concept of IQ ever possessed has broken down very rapidly indeed since the famous (or infamous) paper by Richard Hernstein "I.Q.", Atlantic Monthly, September 1971, discussing race, heredity, and intelligence.

21. Huysmans, "op. cit", (12) (Quoted by permission).
22. Little, J.D.C., "Models and Managers, the Concept of a Decision Calculus", Management Science, vol. 18-8, April 1969.
23. Sander, W.E., "A Scoring Methodology for Assessing the Suitability of Management Science Models", Management Science, vo. 18-10, June 1972. *Evaluation*
24. See Keen, P.G.W., The Implications of Cognitive Style for Individual Decision-Making, unpublished doctoral dissertation, Graduate School of Business Administration, Harvard University, 1973, Chapter 10 passim.
25. Churchman, C.W., "Managerial Acceptance of Scientific Recommendations", California Management Review, 1967.
26. Lewin, K., Field Theory in Social Science, Harper, 1951.
27. Guilford sets even stronger requirements for factor analysis; given his role in developing the technique, his opinion seems especially weighty: Guilford, J.P., Psychometric Methods, McGraw-Hill, 1954, Chapter 16.
28. Siegel, S., Nonparametric Statistics for the Social Sciences, McGraw-Hill, 1956.
29. See L. Schatzman, L. and Strauss, A.L., Field Research: Strategies for a Natural Sociology, Prentice-Hall, 1973.
30. Eg. Olesen, V.L., and Whittaker, E.W., "Role-makings in Participant Observation: Processes in the Researcher-Actor Relationship", Human Organization 26, pp. 273-81, 1967.
31. Schatzman and Strauss, op. cit, p. 58.
32. See Beckhard, R., Organization Development: Strategies and Models, Addison-Wesley, 1969.

33. Eg. Vestinsky, I., Barth, R.T., and Mitchell, V.F., "A Study of OR/MS Implementation as a Social Change Process", unpublished paper, 1973 and Sorensen, R.E. and Zand, D.E., "Improving the Implementation of OR/MS Models by Applying the Lewin-Schein Theory of Change", unpublished paper, 1973.
34. Schein, E., "Management Development as a Process of Influence", Industrial Management Review, II, No. 2, May 1961.
35. Sorensen and Zand, "op.cit." (33)
36. Schein, "op.cit.", p. 161.
37. Forcese, D.P., and Richer, S., Social Research Methods, Prentice-Hall, 1973, pp. 102-3
38. Webb, E. et al, Unobtrusive Measures: Non-reactive Research in the Social Sciences, Rand McNally, 1966.
39. Schein, "op.cit." (34). Beckhard, Schein, Bennis and Kolb, among others, have used this focus throughout their research.
40. This example was chosen since each of the three historians is accepted as a major figure in research on that period, and yet their description, assessment, and conclusions differ overwhelmingly.
41. See Hammond, J.S., "The Roles of the Manager and Management Scientist in Successful Implementation", Sloan Management Review, 15, No. 2, 1974, pp.1-24.
42. Gibson, op.cit., p. 4.
43. Gibson, op.cit., p. 5
44. Sorensen and Zand, "op.cit.", (33).

45. Kuhn, op.cit., (5).
46. Deutsch, K.W., Recent Trends in Research Methods in Political Science, in Charlesworth J.C., A Design for Political Science, 1966, p. 158.
47. Merton, R.K., Social Theory and Social Structure, Free Press, 1957.
48. LaPalambra, J., Parsimony and Empiricism in Comparative Politics: An Anti-Scholastic View in Holt and Turner, op.cit. (19) pp. 133-34.

MMI 16 10

MIT LIBRARIES

3 9080 00367 1523

700-74

MIT LIBRARIES

3 9080 00367 1556

701-74

MIT LIBRARIES

3 9080 00367 1572

702-74

MIT LIBRARIES

3 9080 00367 1598

703-74

MIT LIBRARIES

3 9080 00370 2534

704-74

MIT LIBRARIES

3 9080 00370 2559

705-74

MIT LIBRARIES

3 9080 00370 2575

706-74

